



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Investigation of Maximum Possible OPF Problem Decomposition Degree for Decentralized Energy Markets

Citation for published version:

Loukarakis, E, Bialek, J & Dent, C 2015, 'Investigation of Maximum Possible OPF Problem Decomposition Degree for Decentralized Energy Markets' IEEE Transactions on Power Systems, vol. 30, no. 5, pp. 2566-2578. DOI: 10.1109/TPWRS.2014.2365959

Digital Object Identifier (DOI):

[10.1109/TPWRS.2014.2365959](https://doi.org/10.1109/TPWRS.2014.2365959)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IEEE Transactions on Power Systems

Publisher Rights Statement:

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Investigation of Maximum Possible OPF Problem Decomposition Degree for Decentralized Energy Markets

E. Loukarakis, *Student Member, IEEE*, J. W. Bialek, *Fellow, IEEE* and Chris J. Dent, *Member, IEEE*

Abstract—The need for improved utilization of existing system assets and energy sources, as well as the smooth incorporation of new technologies (such as electric vehicles) into the grid, has prompted the participation of small power consumers and generators in the energy markets. A problem of such scale however cannot be managed in a centralized manner in its full detail. This paper examines the idea of a decentralized approach in clearing the energy market. A general framework for the problem decomposition and its distributed solution is presented and analyzed. A key point of interest in this work is the fundamental question of how far decomposition may be pursued for a given system, while still achieving reasonable convergence properties. The corresponding optimization problem is formulated and solved through a parallel implementation of the Alternating Direction Method of Multipliers (*ADMM*). A thorough investigation of its convergence properties is conducted, and through its coordination with an additional proximal based decomposition method we improve its scalability characteristics.

Index Terms—Energy Markets, Optimal Power Flow, Distributed Optimization, Multi-Agent Systems

I. INTRODUCTION

POWER systems are gradually transitioning from an era of a limited number of large generators and largely inelastic demand, to an era of highly dispersed small scale generation, deferrable demand (through advances in smart metering), large numbers of so-called prosumers (users alternately appearing as generators or consumers) and distributed storage (both dedicated grid-connected and mobile provided by electric vehicles). In order to make full use of the potential benefits this increased granularity and diversity can offer in terms of market economic efficiency, the particular constraints and objectives of each individual have to be taken into account in the market clearing optimization problem. This however would result in a problem of a particularly large scale. Its centralized solution, due to both computational and communications requirements, might not be possible [1]. As a result a suitably designed decentralized approach, if tenable, would be

of particular interest, as it could enable handling the increased size of the problem in an efficient way. In addition, it could allow for a more detailed system representation and the smooth incorporation of demand response services, thus limiting any related inefficiencies [2] or price volatility [3]. Resilience (i.e. ability to work even with loss of some of its components), limited information sharing (especially of financial nature), and fast solution speed (i.e. fast enough to clear the market at a desired time resolution, e.g. every 15min) are three important desirable properties of this decentralized scheme.

A variety of mathematical techniques has been previously used in distributed and / or decentralized approaches to power system control and operation. In [4] a Lagrangian Relaxation (*LR*) with a subgradient method is used. However, the fact that in certain cases this technique may result in oscillating behavior during convergence [5] has prompted the use of a variety of heuristics or augmented Lagrangian approaches. An example of the former may be found in [1] which utilizes a centralized scheme which arbitrarily limits market players demand response capability, or in [6] where the full *AC OPF* is solved using a centralized Newton method to update Lagrange multipliers. In the case of augmented Lagrangian methods, a popular approach is based on the so called auxiliary problem principle (*APP*) [7]. The method's details and general implementation background are discussed in [8, 9, 10]. The results in [11] however indicate that for a poor parameter selection the method may fail to converge. In [12] a proximal point based method (*PPM*) is used which bears close resemblance to the simpler alternating direction method of multipliers (*ADMM*). Reference [13] introduces a serial implementation of the latter and compares it with the previous two methods. The results do not indicate any significant differences. Reference [14] applies the *ADMM* method to randomly generated large scale demand management problems using *DC* load flow, indicating potentially good scalability for the method. Reference [15] extends the method's application to *AC* load flow with test cases mostly focusing on distribution, while [16] provides a more detailed mathematical foundation for the method's use in *OPF* problems and some results in meshed networks of up to 118 buses. In [17, 18, 19] an approach based on the Karush-Kuhn-Tucker optimality conditions decomposition (*OCD*) is used. Even though [20] indicates that this approach might reach a solution faster than Lagrangian based methods, its convergence condition might not be always easy to verify, nor carry-

This work was supported by the UK Engineering and Physical Science Research Council, as part of the Autonomic Power System project under the research grant EP/I031650/1.

E. Loukarakis and C. Dent are with the School of Engineering & Computing Sciences, Durham University, UK (e-mail: {emmanouil.loukarakis, chris.dent}@durham.ac.uk).

J. Bialek is with the Skolkovo Institute of Science and Technology, Moscow, Russia (e-mail: j.bialek@skoltech.ru).

ing out the subsequently required conditioning of the problem in a decentralized manner.

The literature discussed above is significant yet the performance of the aforementioned methods has not been tested for large degrees of decomposition (i.e. large number of optimization subproblems, even down to the individual bus level) with the inclusion of the full AC equations describing the transmission system constraints. Such an analysis is the main task of this paper. Generally, this work proposes and investigates the feasibility of an approach for distributed optimization, for the purpose of clearing the energy market in a decentralized manner. The basic point of interest is the fundamental question of how far decentralization can be pushed within an energy market. In other words, can the *OPF* solution be decomposed down to the individual node, and even down to the individual energy user connected at each node? There is a wealth of literature on distributed *OPF* but, to our knowledge, none before has investigated this particular question. This question has very important implications as, if the answer is positive, then there is a potential that individual market players could trade directly with each other without relying on the System Operator (SO). In our simulations we concentrate on transmission networks but the proposed method is general and the results relate to both transmission and distribution.

A key point for decentralized energy markets is that they should not require any confidential financial information to be centrally reported (e.g. costs / profits). For this reason and also due to its robust convergence characteristics and limited number of parameters, we have employed the *ADMM* method. We assume a simplified but reasonable market model, but by no means do we offer a complete solution that covers all aspects of trading and market clearing. Instead we focus on fundamental convergence considerations of the decentralized approach and provide a general framework intended for use in energy market operations into which additional details can be incorporated. More specifically the main contributions of this work are:

- We have investigated through extensive simulations the *ADMM* method's convergence performance in the *OPF* problem. Our test systems include the *IEEE* 24, 57, 118 and 300 bus systems, and a 707 bus system based on the UK network. The *ADMM* method has been applied before in power systems, but compared to [21] we focus in non-convex *OPF* formulations in meshed networks. Compared to [15, 16] we use a problem reformulation introduced in [22] extended to different decomposition structures with respect to demand disaggregation. In addition we relate the method's parameters to characteristics of the *OPF* problem and identify suitable settings to improve convergence in typically hard to converge network congested cases.
- Through the introduction of aggregators and a variable penalty proximal decomposition algorithm for the corresponding subproblems we illustrate the value of aggregation from an optimization perspective and improve the scalability of the initial *ADMM* decomposition scheme. The idea of coordinating different distributed optimization algorithms in order to improve convergence is a novel part of this work. In

addition compared to other approaches that consider demand decomposition (e.g. [23]) the proposed scheme does not require transmitting any financially sensitive information (e.g. objective function costs).

- We investigate various degrees of decomposition of the network, down to the individual bus level, combined with decomposition of demand and generation connected at a given bus into independent blocks. The results provide significant insight in the scalability of the proposed approach both with respect to network and user decomposition, and may serve as a reference for future testing of relevant distributed optimization applications.

This paper is organized as follows: Section II describes the centralized problem equations, section III provides the mathematical background behind decomposition and the subproblems structure, while section IV presents and analyses simulation results. Finally section V summarizes the conclusions of this work. With respect to mathematical notation: we use bold font for vectors or matrices (e.g. \mathbf{z}) and italics for scalars (e.g. z); $\mathbf{z}_{(x,y)}$ indicates element (x, y) of matrix \mathbf{z} ; $\text{diag}\{\mathbf{z}\}$ indicates a diagonal matrix whose diagonal elements are the elements of vector \mathbf{z} ; the operator $\|\cdot\|_2^2$ denotes the squared Euclidean norm.

II. OPTIMIZATION PROBLEM FORMULATION

The target in clearing the market is maximization of social benefit, i.e. the minimization of costs (negative utility) [6, 24]:

$$\min \left\{ \sum_{i \in [1..n_u]} \overbrace{(u_i(\mathbf{s}_{(i)}))}^{f_0} \right\} \quad (1)$$

Where:

- u_i Cost function of the i -th client.
- n_u The number of clients / network users.
- \mathbf{S} Clients apparent power input / output $n_u \times 1$ vector. This may further be written as $\mathbf{S} = \mathbf{P} + j\mathbf{Q}$, where \mathbf{P} and \mathbf{Q} the active and reactive power components respectively.

This optimization problem is subject to multiple constraints. First and foremost these involve the transmission network constraints:

$$\mathbf{c}_u \mathbf{S} = \text{diag}\{\mathbf{V}\}(\mathbf{YV})^* \quad (2)$$

$$\underline{\mathbf{V}} \leq |\mathbf{V}| \leq \bar{\mathbf{V}} \quad (3)$$

$$|\text{diag}\{\mathbf{c}_t \mathbf{V}\}(\mathbf{Y}_t \mathbf{V})^*| \leq \bar{\mathbf{T}} \quad (4)$$

Where:

- n_b The number of network buses.
- n_l The number of transmission system lines.
- $\bar{\mathbf{T}}$ Transmission line apparent power limit $n_l \times 1$ vector.
- \mathbf{V} Complex bus voltage $n_b \times 1$ vector. $\bar{\mathbf{V}}$ and $\underline{\mathbf{V}}$ denote the upper and lower bounds on voltage magnitude respectively.
- \mathbf{Y} Bus admittance $n_b \times n_b$ matrix.
- \mathbf{Y}_t Line admittance $n_l \times n_b$ matrix with $\mathbf{Y}_{t(i,k)} = y_k$ if line i starts from bus k , $\mathbf{Y}_{t(i,k)} = -y_k$ if line i ends at bus k , and 0 other wise; y_k is the admittance of line k .
- \mathbf{c}_u Client to bus $n_b \times n_u$ connection matrix with $\mathbf{c}_{u(i,k)} = 1$ if client k is connected to bus i , and 0 otherwise.
- \mathbf{c}_t Line to bus voltage $n_l \times n_b$ connection matrix with $\mathbf{c}_{t(i,k)} = 1$ if line i starts from bus k , and 0 otherwise.

Equation (2) describes the power balance constraints at each

bus, equation (3) the voltage magnitude constraints and equation (4) the transmission capacity constraints (a similar set of equations is written for the power at the end of the line). Individual network users are described by the following simplified generic equations:

$$u_i = c_{2i} \mathbf{P}_{(i)}^2 + c_{1i} \mathbf{P}_{(i)} \quad (5)$$

$$\underline{\mathbf{P}}_{(i)} \leq \mathbf{P}_{(i)} \leq \overline{\mathbf{P}}_{(i)} \quad (6)$$

$$\underline{\mathbf{Q}}_{(i)} \leq \mathbf{Q}_{(i)} \leq \overline{\mathbf{Q}}_{(i)} \text{ or } \mathbf{Q}_{(i)} = f_Q(\mathbf{P}_{(i)}) \quad (7)$$

Where $i \in [1, \dots, n_u]$ and:

c_2, c_1 Variable costs coefficients for generators. For demand $c_2 = 0, c_1 > 0$ and equal to the value a client associates with energy use. It can be thought of as an equivalent to the value of lost load (*VOLL*) which is assumed to be about 100 times the value of energy at peak demand.

f_Q Function of reactive power as a function of active power. For e.g. demand operating at a fixed load factor this is simply a linear function. For devices where reactive power is independent from active power, this function does not apply and only reactive power limitations are taken into account.

Equations (5)-(7) imply that all demand may be curtailed. In this way, even in cases where demand may not be fully covered due to lack of adequate generation or transmission capacity, the overall optimization problem is still feasible. As shown in Fig.1, the right generation characteristic curve intersects with the vertical part of the demand curve setting the price at marginal generation cost, while the left generation curve is inadequate to cover all demand.

III. OPTIMIZATION PROBLEM DECOMPOSITION

This section addresses two basic questions: a) how to determine the subproblems into which the initial optimization problem will be decomposed to, and b) how to perform the decomposition assuming the desired subproblems are known. Each subproblem is considered to be managed by an agent, i.e. an entity which handles all necessary communications and runs the required optimization routines.

A. Decomposition Mathematics

As a first step to bring the problem into a suitable form for decomposition, one fictitious bus is introduced at the middle of each line that connects systems which are managed by separate agents. In a similar fashion a fictitious node may be introduced for each demand or generation block that we want to handle as a separate subproblem. The variables associated with that particular bus / node are duplicated as seen on Fig.2.

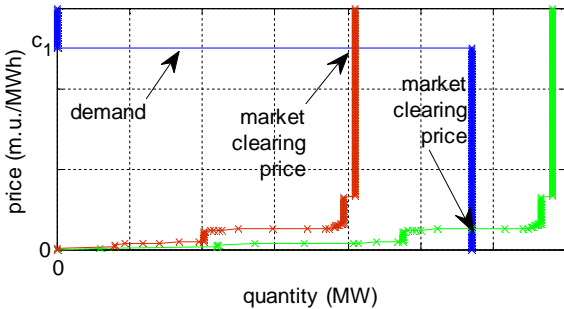


Fig.1. Demand curve and two example generation curves. When generation is inadequate the market clearing price is c_1 or equal to *VOLL*.

Let $\mathbf{U}_e = [\mathbf{S}_e; \mathbf{V}_e]$ be the duplicated variables vector, $\mathbf{h}_c \mathbf{U}_e = 0$ the linear non-separable constraints ensuring duplicated variables equality, $\mathbf{U} = [\mathbf{S}; \mathbf{V}; \mathbf{U}_e]$ a vector of all optimization variables, and $\mathbf{h}_e(\mathbf{U}) \leq 0$ the extended non-linear separable equations, similar in form to (2)-(4) and (5)-(7), that take into account the duplicated variables. The coupling constraints $\mathbf{h}_c \mathbf{U}_e$ are now handled using the *ADMM* method [25]. First the initial optimization problem is reformulated as follows:

$$\min_{\mathbf{z}, \mathbf{U}} \{f_o(\mathbf{U}) + g(\mathbf{z}) : \mathbf{z} = \mathbf{U}_e, \mathbf{h}_e(\mathbf{U}) \leq 0\} \quad (8)$$

$$g(\mathbf{z}) = \begin{cases} 0 & \text{if } \mathbf{h}_c \mathbf{z} = 0 \\ +\infty & \text{if } \mathbf{h}_c \mathbf{z} \neq 0 \end{cases} \quad (9)$$

The augmented Lagrangian is:

$$\mathcal{L}_\rho(\mathbf{U}, \mathbf{z}, \boldsymbol{\lambda}_e) = f_o(\mathbf{U}) + g(\mathbf{z}) + \boldsymbol{\lambda}_e^T (\mathbf{U}_e - \mathbf{z}) + (\rho/2) \|\mathbf{U}_e - \mathbf{z}\|_2^2 \quad (10)$$

where ρ is the penalty factor and $\boldsymbol{\lambda}_e$ a vector of Lagrange multipliers corresponding to the constraints $\mathbf{z} = \mathbf{U}_e$. Then starting from an estimate of $\boldsymbol{\lambda}_e$ and \mathbf{z} the following steps are repeated until convergence:

$$\mathbf{U}^{k+1} = \underset{\mathbf{U}}{\operatorname{argmin}} \{ \mathcal{L}_\rho(\mathbf{U}, \mathbf{z}^k, \boldsymbol{\lambda}_e^k) : \mathbf{h}_e(\mathbf{U}) \leq 0 \} \quad (11)$$

$$\mathbf{z}^{k+1} = \underset{\mathbf{z}}{\operatorname{argmin}} \{ \mathcal{L}_\rho(\mathbf{U}^{k+1}, \mathbf{z}, \boldsymbol{\lambda}_e^k) \} \quad (12)$$

$$\boldsymbol{\lambda}_e^{k+1} = \boldsymbol{\lambda}_e^k + \rho(\mathbf{U}_e^{k+1} - \mathbf{z}^{k+1}) \quad (13)$$

where k is the iteration number. Due to the fact that the last term in (10) is non-separable, instead of performing a minimization simultaneously over \mathbf{U} and \mathbf{z} , minimization is performed separately, in an alternating fashion. A first minimization step in (11) optimizes \mathbf{U} with \mathbf{z} fixed, and the second minimization step in (12) optimizes \mathbf{z} with \mathbf{U} fixed. Equations (11-13) may be rewritten into the following form:

$$\mathbf{U}^{k+1} = \underset{\mathbf{U}}{\operatorname{argmin}} \left\{ \overbrace{f_o(\mathbf{U})}^{\text{cost term}} + \overbrace{(\boldsymbol{\lambda}_e^k)^T \mathbf{U}_e}^{\text{price term}} + \overbrace{(\rho/2) \|\mathbf{U}_e - \mathbf{z}^k\|_2^2 : \mathbf{h}_e(\mathbf{U}) \leq 0}^{\text{ADMM penalty term}} \right\} \quad (14)$$

$$\mathbf{z}^{k+1} = \underset{\mathbf{z}}{\operatorname{argmin}} \{ \|\mathbf{U}_e^{k+1} - \mathbf{z}\|_2^2 : \mathbf{h}_c \mathbf{z} = 0 \} \quad (15)$$

$$\boldsymbol{\lambda}_e^{k+1} = \boldsymbol{\lambda}_e^k + \rho(\mathbf{U}_e^{k+1} - \mathbf{z}^{k+1}) \quad (16)$$

Equations (14)-(16) are separable but require exchange of coupled variable values between directly connected subproblems. The introduction of the auxiliary variable \mathbf{z} is what effectively induced this separability which enables the parallel solution of the generated subproblems. Equation (15) presup-

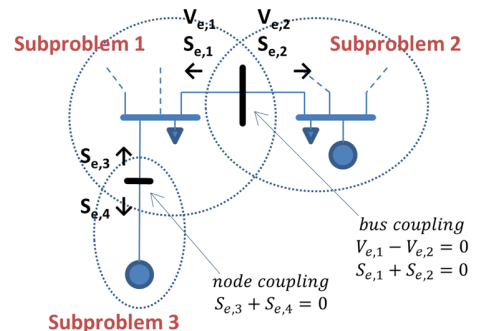


Fig.2. Fictitious buses and nodes, and duplication of variables.

poses that the initial estimate for the Lagrange multiplier of a fictitious bus or node is selected so that $(\lambda_e^0)^T \mathbf{z} = 0 \forall \{\mathbf{z}: \mathbf{h}_e \mathbf{z} = 0\}$. A proof may be found in section VI-B.

Convergence requires that the primal and dual residuals should be sufficiently small. This may be expressed as $|\lambda_e^{k+1} - \lambda_e^k| \leq \varepsilon$ and $\rho \|\mathbf{z}^{k+1} - \mathbf{z}^k\| \leq \varepsilon$. It has been suggested in [25, 14] that a variable penalty factor keeping the two residuals in the same order of magnitude may improve convergence. However depending on the penalty factor update step it was observed that this might not always be the case. Therefore in this work a constant penalty factor was used.

B. Decentralized System Structure

The basic structure of the problems solved in this work is illustrated in Fig.3 while the corresponding algorithmic flowcharts and subsequent information exchange between agents may be seen on Fig.4. Three different decomposition schemes were considered:

- *Scheme A (network decomposition)*: Each subproblem contains a part of the transmission network. The agent managing a subproblem is assumed to have complete knowledge of utility functions and constraints of users connected to his network area. Such an agent may be considered as the equivalent of a transmission system operator and is designated as TSO^* . All TSO^* agents operate in parallel, i.e. for the algorithm to progress to the next iteration all agents have to solve their respective optimization subproblems. This scheme, which is similar to standard approaches in literature for power system areas coordination, is used to test scalability with respect to the number of network subproblems.
- *Scheme B (network and user decomposition)*: The power system is decomposed simultaneously to network areas and individual network user blocks. For each network area a TSO agent manages the corresponding subset of transmission constraints (2)-(4). In contrast with the TSO^* agents the TSO problem no longer contains any user constraints or any cost term in the objective function which is similar in form to (14). A set of network users, represented by constraints similar to (5)-(7), is managed by a microgrid operator (MO). An MO agent would in practice represent any number of nearby located users (at the extremes it could represent a single user or all the users at a specific bus), and would have to deal with the peculiarities of end user equipment and demands (e.g.

communication issues, unexpected requests etc.). Again all agents work in parallel. This scheme is used to test scalability with respect to the disaggregation of network users.

- *Scheme C (network and user decomposition with user aggregation)*: This is a two-step decomposition scheme used to test the effects of aggregation with respect to network users. First the initial problem is decomposed to TSO subproblems (as in scheme B) and bus aggregator subproblems. These bus aggregators could be considered the equivalent of a distribution system operator (DSO). Each bus aggregator subproblem (designated as DSO^*) contains the objectives and constraints of all users located at a specific bus. The DSO^* subproblem is the equivalent of having a single MO managing all the demand at a transmission bus. Following, each DSO^* subproblem is further decomposed to individual MO subproblems and an aggregator subproblem (designated as DSO). The latter does not contain any user constraints and involves only the power balance constraint between the sum of MO s and the transmission grid. The DSO effectively sums up the MO agents response, thus limiting communications and computational requirements for TSO agents. In the following for ease of presentation of the aggregator problem we focus on active power only. The extension to include reactive power is straightforward. The DSO^* subproblems are generally problems of following form:

$$\min_{\mathbf{p}_D \in C_D, P_{eD}} \left\{ \overbrace{f_{0D}(\mathbf{p}_D)}^{\text{cost}} + \overbrace{\lambda_{eD}^k P_{eD}}^{\text{price term}} + \overbrace{(\rho_D/2) \|P_{eD} - z_D^k\|_2^2}^{\text{ADMM penalty term}} \right\} \quad (17)$$

Where C_D is the intersection of multiple user constraints sets of the form (5)-(7) and in addition:

- P_{eD} A single element of \mathbf{U}_e which is associated with the active power of the DSO as seen from the transmission level.
- \mathbf{p}_D A vector consisting of elements of \mathbf{U} which correspond to the active power of MO s associated with the DSO .
- λ_{eD}^k A single element of λ_e^k associated with the active power of the DSO .
- f_{0D} The part of the original objective function f_0 that corresponds to the aggregator based on his associated users.
- ρ_D, z_D^k Penalty factor and auxiliary variable of the k -th iteration of the original ADMM decomposition associated with the DSO .

The adaptive proximal decomposition method ($APDM$) used for this problem's solution involves the following steps:

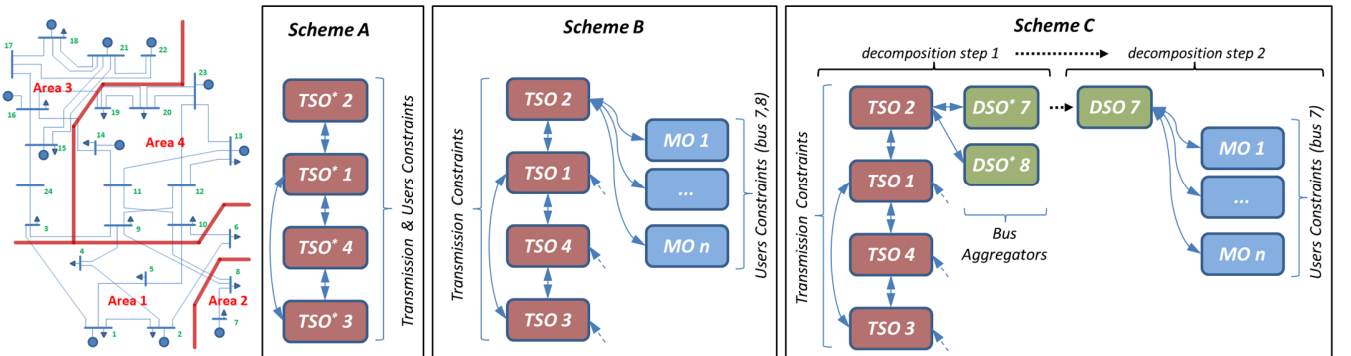


Fig.3. Tested decomposition schemes and sample illustration of generated subproblems interrelations. Each block represents a subproblem / agent and each line indicates subproblem (physical) coupling and a required bidirectional communications link.

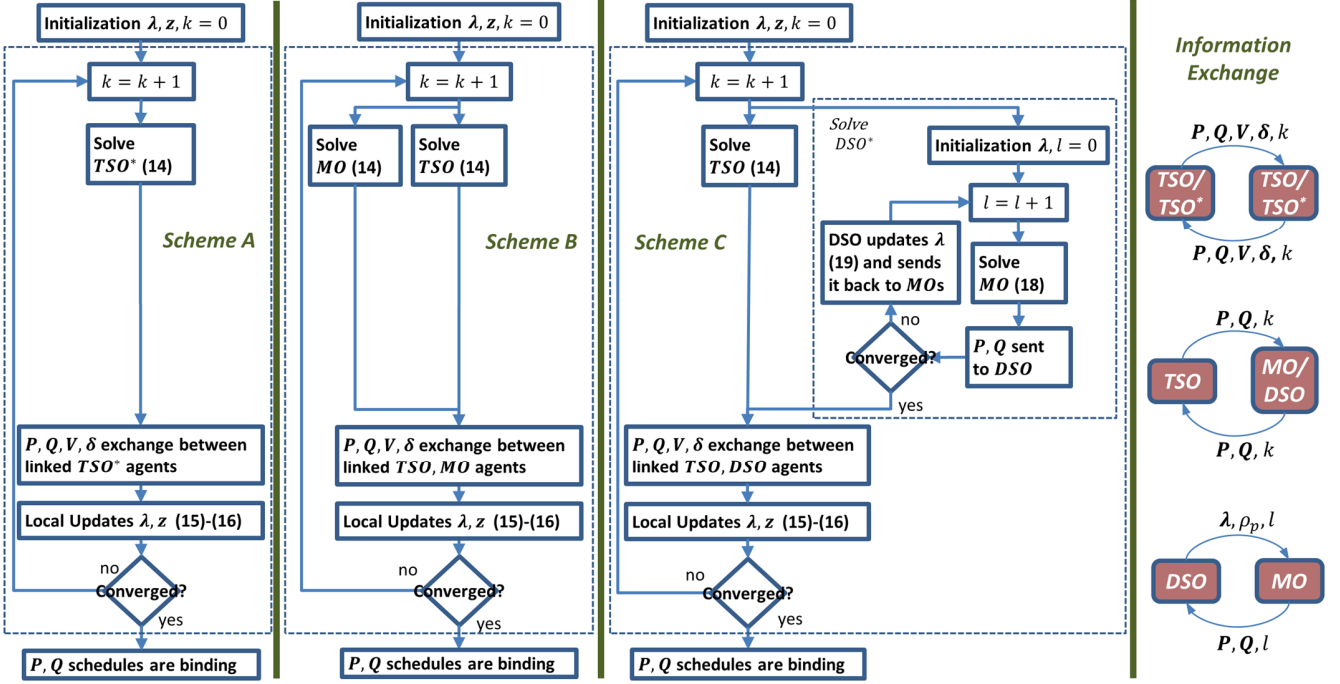


Fig.4. Flowcharts for the decentralized solution for each of the decomposition schemes and indicative illustration of information exchange between different types of agents within a single iteration. Initialization would typically use the values of λ_e, z of the last algorithm run. Regarding the information exchanges the iteration count is passed in order to facilitate agent synchronization. For purposes of error checking λ_e values could also be periodically transmitted. It should be noted that all schemes use synchronous implementation of the methods where at each iteration all subproblems have to be solved and relevant information collected, before progressing to the next.

$$\mathbf{P}_D^{l+1} = \underset{\mathbf{P}_D \in \mathcal{C}_D}{\operatorname{argmin}} \left\{ \underbrace{\frac{\text{cost}}{f_{0D}(\mathbf{P}_D)}}_{\text{cost}} + \underbrace{\lambda_D^l \mathbf{1}_D \mathbf{P}_D}_{\text{price term}} + \underbrace{\left(\frac{\rho_p}{2} \right) \|\mathbf{P}_D - \mathbf{P}_D^l\|_2^2}_{\text{APDM penalty term}} \right\} \quad (18)$$

$$\lambda_D^{l+1} = \nabla_{P_{eD}} \left(\lambda_{eD}^k P_{eD} + \frac{\rho_D}{2} \|P_{eD} - z_D^k\|_2^2 \right) \Big|_{P_{eD} = \mathbf{1}_D \mathbf{P}_D^{l+1}} \quad (19)$$

$$= \lambda_{eD}^k + \rho_D (\mathbf{1}_D \mathbf{P}_D^{l+1} - z_D^k)$$

Where l is the iteration count. In addition:

- ρ_p Penalty factor for APDM method.
- λ_D Marginal price associated with the constraint $P_{eD} = \mathbf{1}_D \mathbf{P}_D$ (power balance at the DSO node).
- $\mathbf{1}_D$ A row vector of ones with length equal to the number of MOs associated with the DSO (i.e. length of \mathbf{P}_D).

Equation (18) is separable with respect to MOs and represents the MO subproblems, while equation (19) constitutes the DSO (price update) subproblem. Convergence is achieved if $\max\{\|\mathbf{P}_D^{l+1} - \mathbf{P}_D^l\|\} \leq \varepsilon$. Assuming accuracy in the order of 10kW is desirable, ε should be set accordingly. For a suitably selected ρ_p value, power converges following a quickly damped oscillation around its optimal value. However it is not easy to estimate the correct penalty value and too high a value will delay convergence. A simple way to achieve good performance is by setting initial bounds $\underline{\rho_p} = 0, \overline{\rho_p} = \rho_D n_a$ (where n_a is the number of MOs managed by the DSO) and using the following empirical updating scheme every few iterations:

- If $\max\left\{\frac{(|\mathbf{P}_D^{l+1} - \mathbf{P}_D^l| - |\mathbf{P}_D^l - \mathbf{P}_D^{l-1}|)}{|\mathbf{P}_D^{l+1} - \mathbf{P}_D^l|}\right\} \leq \frac{1}{2} \rightarrow \underline{\rho_p} = \rho_p$. The inequality if valid implies large oscillations around the optimum and consequently slow convergence due to a low penalty value. Thus the lower bound is increased.
- If $\max\{(\mathbf{P}_D^{l+1} - \mathbf{P}_D^l)(\mathbf{P}_D^l - \mathbf{P}_D^{l-1})\} \geq 0 \rightarrow \overline{\rho_p} = \rho_p$. The inequality if valid implies that convergence is slow due to a high penalty factor value. Thus the upper bound is decreased.
- Set $\rho_p = (\overline{\rho_p} + \underline{\rho_p})/2$.

The reason for using this more complex proximal scheme instead of a simple price-based approach is that users with non-strongly convex (e.g. linear) utility curves may be effectively managed. For MOs with identical utility functions and a fixed penalty factor this method is equivalent to the one presented in [26].

It should be emphasized that all the proposed decomposition schemes have the important advantage that the agents handling each subproblem do not have to disclose any confidential economic information to other agents. The coordination is done by exchanging information only about the power quantities each agent is willing to trade at the specific price (or Lagrange multiplier) estimate of each iteration. Constraints and objectives are managed locally by each agent and consequently privacy over costs and limitations is retained. Furthermore the price updates are performed independently at each node, thus no centralized control is required. The overall decomposition process and structure of the subproblems are also further clarified in the example of section VI.

C. Network Partitioning

Determination of optimization subproblems at the transmission level presupposes partitioning the network which is a large integer programming problem, typically solved through various heuristic algorithms. In this work we use a spectral clustering algorithm [27]. Assuming a bus adjacency matrix is defined based on electrical distance [28] the bus admittance matrix \mathbf{Y} (neglecting shunt admittances) is the associated Laplacian. If k partitions are required the following steps are involved:

1. Calculate the normalized Laplacian $\mathbf{L}_n = \mathbf{D}^{-0.5} \mathbf{Y} \mathbf{D}^{-0.5}$, where $\mathbf{D} = \text{diag}(\mathbf{Y}_{(1,1)}, \dots, \mathbf{Y}_{(n_b, n_b)})$.
2. Find the eigenvectors v_1, \dots, v_k corresponding to the k smallest eigenvalues of \mathbf{L}_n and form the matrix $\mathbf{V} = [v_1, \dots, v_k]$.
3. Determine the matrix \mathbf{V}_n by normalizing each of \mathbf{V} 's rows to have unit length.
4. Treating each row of \mathbf{V}_n as a point in \mathbb{R}^k , cluster them using a k -means algorithm.
5. Assign a bus i into partition j if row i of \mathbf{V}_n was assigned to cluster j .

Spectral clustering is a commonly used partitioning method in power systems [29] and as such is reasonably adequate for illustrating the impact of increasing network decomposition. Of course this does not necessarily mean that this is the optimal way to partition the system, and other methods, e.g. spectral clustering variants [30] or based on fuzzy or evolutionary algorithms [31, 32], could have been used instead. Identifying an optimal partitioning would require appropriately linking the characteristics of currently available partitioning methods with those of the distributed optimization algorithm, and suitably extending and investigating their performance. This is a considerable task far beyond the purposes of the present work.

D. Practical Implementation Considerations

The proposed distributed optimization methods are generally robust to changes in the network and will continue working even if users enter or leave the optimization process, or part of the network is lost due to an outage. However in order to achieve convergence a fixed problem structure is required. This presupposes rational agent behavior (i.e. consistent bidding / response to received prices) and reliable communications. Regarding agent behavior, in this work it is assumed that each agent is equipped with a digital device that solves a generic predetermined form of optimization problem and handles the necessary communications. The parameters of that problem would be provided by the agent and while they could be changed, they would remain 'locked' during a distributed optimization run. Implemented in such a way, in terms of agent behavior, the distributed approach would not face any different problems than a centralized solution would. At the same time the response speed of an agent is as fast as its processing power and communications infrastructure allow. The overall behavior of an agent (i.e. how he determines the parameters of his optimization problem in terms of utility and constraints) between consecutive optimization runs is still an interesting research subject. A wide range of relevant work

and ideas may be found in papers related to bidding practices in energy markets (e.g. [33, 34]), or agent-based operations modeling (e.g. [35]).

On the other hand regarding communications, recurring errors in the transmission of information could delay convergence significantly, while persistent errors for an agent could be perceived as absence of that agent from the market and could lead to a wrong optimization solution. Reliable communications are of paramount importance for any decentralized scheme but how these would be achieved is outside the scope of this work. We should point out however that with respect to the proposed decomposition schemes, high reliability communications would be required between *TSOs* as significant amounts of energy may be expected to be traded by them (e.g. on the order of several MW to GW). The required communications reliability for *DSOs* (typically on the order of a few to several MW) could be more relaxed depending on their size, while for *MOs* a high level of reliability might be too expensive to achieve. After all given their smaller size (e.g. on the order of several kW to MW) unreliable information from a few *MO* agents in most cases would not significantly affect the solution. Decomposition scheme C can help in such cases as the *DSO* could approximate an *MO* solution in case of a detected communications fault. This approximation could also be possible for both schemes B and C between *TSOs* and *DSOs/MOs* (e.g. with the *TSOs* simply using demand forecasts). Overall the selection of suitable communications methods and protocols for each decomposition level is a subject of great importance, and the answer regarding what is most efficient in terms of performance and costs is not yet clear. Relevant information for the interested reader may be found in e.g. [36, 37].

Closing it should be noted that in practice it might be the case that different market players are subjected to different market rules. These rules are related with the way market players perceive their utility and in certain cases, some of their constraints. The overall decomposition approach is independent of that, and can incorporate any such market variations.

IV. INDICATIVE RESULTS

The test systems used in this work include the 24, 57, 118, 300 bus IEEE test systems and a 707 bus representation of the UK network. Both a full AC and a non-linear DC formulation (by setting voltages to unity and neglecting reactive power equations) are investigated. No reactive power data were available for the UK test system and as a result it was tested only with the non-linear DC formulation.

A. General Observations (scheme A)

The applied distributed optimization method has effectively two parameters which affect convergence. The first is the penalty factor ρ and the second the tolerance ε . Fig.5 illustrates the general convergence progress of the method for the 24 bus IEEE RTS system at peak demand. The problem is decomposed using scheme A to 24 subproblems (i.e. down to individual bus level) using the full AC load flow equations with a flat start (i.e. $\lambda_e^0 = 0, \mathbf{z}^0 = 0$). Fig.5 shows that about 200 itera-

tions are required for a tolerance of 10^{-2} but about twice as many are needed if tolerance is set to 10^{-3} . For $\varepsilon = 10^{-1}$ the maximum error in marginal prices compared to centralized *OPF* solution was 3.8%. For $\varepsilon \leq 10^{-2}$ the error was less than 1%. It should be noted that the tolerance is not directly related to accuracy in marginal prices, so setting a high value on tolerance (e.g. higher than 10^{-2}) might give inaccurate results in certain cases. Consequently, despite the potential high increase in iterations, a low tolerance (i.e. on the order of 10^{-3}) is advisable.

Regarding the effect of ρ on convergence it may be seen from (14) that for very low values the method nearly degenerates into basic Lagrangian Relaxation which implies that for non-strongly-convex objective functions may fail to converge. On the other hand high values of ρ typically result in highly oscillatory behavior and subsequently delayed convergence. Intuitively the convergence performance would be dependent on the interaction of the last two terms in (14) and consequently on the value of λ_e^*/ρ . Fig.6 illustrates convergence performance for a large variety of cases which involve all our test systems at various degrees of decomposition (up to the individual node level) and loading conditions (including cases which require demand curtailments), as a function of λ_e^*/ρ , where λ_e^* is the maximum Lagrangian multiplier value at the optimization problem solution. As may be seen the number of iterations in all cases is minimized for $\lambda_e^*/\rho \approx 6-8$. Thus ρ should be accordingly set. In practice, a good estimate of λ_e^* would generally be available based on forecasts or any forward market solutions. In cases where this estimate is far from the actually realized prices, simple logical rules could be used to periodically adjust the penalty factor during the algorithm execution.

In Fig.7 convergence results are presented for a modified version of the 24 bus *IEEE RTS*, where the capacity in certain transmission lines was reduced, resulting in a congested state with highly divergent prices. AC load flow with a flat start (i.e. $\lambda_e^0 = 0, \mathbf{z}^0 = 0$) and decomposition to the nodal level was used in the simulations. As may be seen the AC formulation leads to an oscillatory behavior around the optimal marginal prices (i.e. the prices a centralized solution yields) and has difficulty in converging. The reason was the interactions between the reactive and active power coupling variables. More specifically, within a subproblem, the local solution would manipulate voltage amplitudes / reactive power in order to enable procuring active power at a reduced cost, given that the penalty factor for them all is the same. However that would imply large voltage / reactive power deviations over the next *ADMM* iteration. This is a cycle that is constantly repeated with the marginal energy prices fluctuating around their optimal value. This shortcoming of the original form of the method can be resolved by modifying the Lagrangian penalty terms from $(\rho/2)\|\mathbf{U}_e - \mathbf{z}\|_2^2$ to $(1/2)(\mathbf{U}_e - \mathbf{z})\boldsymbol{\rho}(\mathbf{U}_e - \mathbf{z})^T$. The matrix $\boldsymbol{\rho}$ has to be set so that the penalty factors associated with voltage magnitudes and reactive power are an order higher than those associated with voltage angles and active power. Intuitively this amounts to solving subproblems where from an active power variables viewpoint, duplicated reactive power

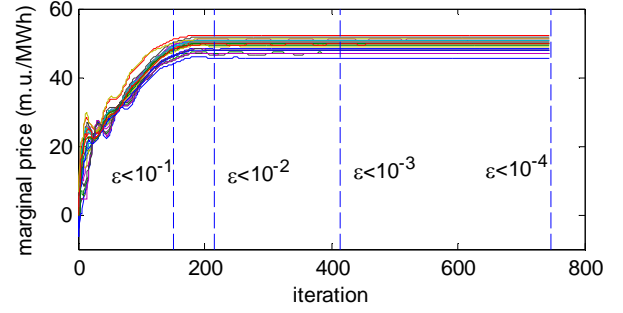


Fig.5. Marginal prices (in monetary units per *MWh*) convergence for *IEEE RTS* decomposed to individual bus level. The vertical lines indicate the iteration when a specific convergence tolerance is reached.

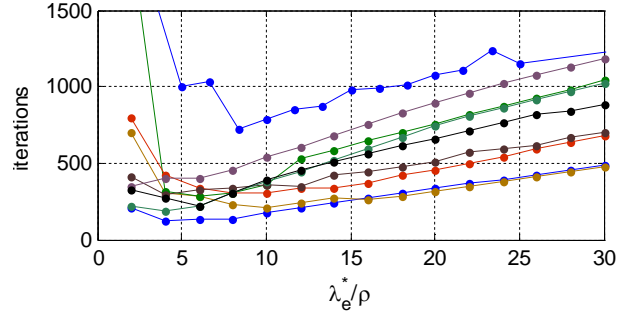


Fig.6. Effect of penalty factor ρ on iterations required for convergence with a tolerance $\varepsilon = 10^{-2}$ for a variety of cases for both AC and non-linear DC formulations.

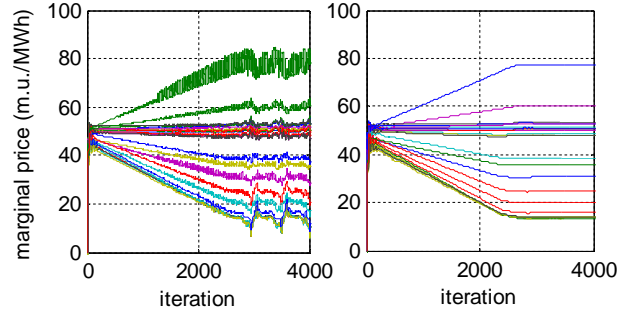


Fig.7. Convergence for the modified *IEEE RTS* high congestion case with full AC equations, decomposition to individual bus level and uniform penalty factors (left), non-uniform penalty factors (right).

related variables are fixed. In this case convergence to the optimal marginal prices is achieved. It should be noted that the non-linear *DC* simulation converges without difficulties.

Summing up, the observations made in this subsection provide general guidelines to setting the method's parameters at any given optimization case, prior to running the case itself. Using the non-uniform penalty factors for AC problem formulations, and associating penalty factors with the expected marginal prices of the system are advisable for achieving good convergence speed.

B. Network-wise Scalability (scheme A)

This subsection investigates the decomposition algorithm's scalability with respect to the number of network subproblems, using decomposition scheme A. Each test system was decomposed to an increasing number of subproblems. Fig.8

shows the results when using the full *AC* equations, while the results with nonlinear *DC* equations are presented in Fig.9. The penalty factor ρ was set for each system based on the guidelines of section IV-A, while the convergence tolerance was set to $\varepsilon = 10^{-3}$. As may be expected as the number of agents / subproblems increases, the number of iterations to convergence also tends to increase. It is of note that each system scales differently, e.g. the 24 or 57 bus systems on average have a much steeper increase in iterations than the 118 or 300 bus systems. However a larger system does not necessarily imply a more difficult convergence (e.g. convergence for the 118 bus system can be faster than convergence for the 24 or 57 bus system). Performance in terms of iterations appears to be more case specific, rather than system size specific.

It may be observed that for the *UK* network after 4 partitions there is a steep increase in iterations. From 5 to 11 partitions the iterations number does not significantly vary, while for more than 11 partitions the increase in iterations is quite significant. The peculiarity of this system, which the other test systems do not share, is that due to a small number of active transmission constraints there is an increased energy price in some buses, significantly higher than the price in other nodes. For a given system partitioning, the effect of those constraints on price might not propagate quickly enough through the network, or in other words the Lagrangian multipliers update can be very slow. Delayed convergence in congested cases especially as the decomposition degree increases may be an issue.

It is interesting to note that often the *AC* formulation converged faster than the nonlinear *DC* despite the fact that the problem in the latter case is essentially much simpler. Consequently the size and complexity of the subproblems does not seem to be a defining factor as far as convergence speed is concerned. Furthermore, for a given system, as the number of subproblems increases, while the general trend is an increase

in iterations, the presented curves show a fluctuating behavior (e.g. for *AC* load flow and 90 subproblems the 300 bus system converges slower than when it is decomposed to 105 subproblems). This implies that the way the system is partitioned can greatly affect convergence.

Overall the results in this section indicate that the convergence speed of the method depends more on the case under study and the extreme conditions (e.g. congestion) that a system may be subjected to, than the size of the problem and number of constraints. When investigating the applicability of *ADMM* in a power system, typical congested scenarios should be studied and special consideration should be given to the method used for network partitioning. Regarding the interpretation of the results already presented here, e.g. for the *UK* network, assuming a maximum computational plus communications latency time of 1s for each subproblem iteration and a market clearing frequency of about 15min, it is not possible to partition the system to more than 4 areas using this spectral clustering approach, as the required time for convergence would be more than the actually available time. It should be noted however that warm-starting the algorithm (e.g. based on forecasts) can improve these results.

C. User-wise Scalability (schemes B & C)

An important property of any distributed scheme is its ability to manage sufficiently fast a large number of network users. This section investigates how the proposed schemes perform for a given network structure as granularity on the demand side increases, i.e. when different clients connected to the same node are handled increasingly individually, rather than as a single aggregate client. Tests are based on decomposition schemes B and C. For any single subproblem the total time for a single iteration would be $t_1 = t_c + t_l$, i.e. the sum of the subproblem local solution time plus the communica-

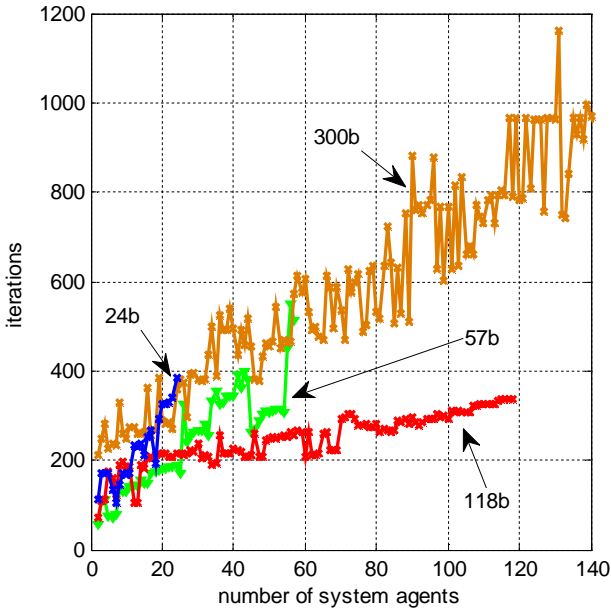


Fig.8. Iterations to convergence as a function of the number of *TSO** agents (i.e. number of areas) using the full *AC* equations.

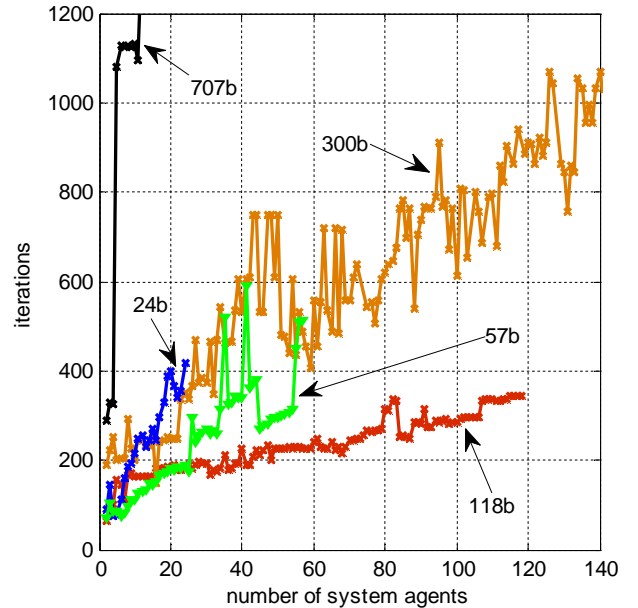


Fig.9. Iterations to convergence as a function of the number of *TSO** agents (i.e. number of areas) using the nonlinear *DC* equations.

tions latency time (time required for sending the information to another agent). The total execution time for scheme B and C may be given by the following relations:

$$t_B = \sum(\max\{t_{1(TSO_1)}, t_{1(TSO_2)}, \dots, t_{1(MO_1)}, t_{1(MO_2)}, \dots\}) \quad (20)$$

$$t_C = \sum(\max\{t_{1(TSO_1)}, t_{1(TSO_2)}, \dots, t_{1(DSO_1^*)}, t_{1(DSO_2^*)}, \dots\}) \quad (21)$$

$$t_{1(DSO^*)} = \sum(\max\{t_{1(MO_1)}, t_{1(MO_2)}, \dots\} + t_{1(DSO)}) \quad (22)$$

The operator Σ denotes summation over all iterations. Currently there is no fully fledged communications standard for the smart grid, and as such it is difficult to predict latency values. Therefore in the examples that follow we assume for all subproblems $t_l = 0.1s$. It should be noted that this is an admittedly simple model for latency; however it is considered adequate for the purpose of comparing the performance of the different decomposition schemes. A more realistic modeling of latency times is an important open problem requiring both new theoretical derivations and practical experimentation.

In the following paragraphs we investigate the comparative performance of the two decomposition schemes assuming a single *TSO* agent. Our test results are illustrated in Fig.10 and involve the disaggregation of demand to an increasing number of *MO* problems by randomly breaking down the initial demand blocks. Parameters for the decomposition algorithms were set based on the guidelines of the previous sections.

- *Case 1 (IEEE RTS 24bus system – base data)*: In this case the *TSO* subproblem is small in size and solved fast. The solution of *DSO*^{*} subproblems takes typically longer. For a fixed penalty factor value, iterations for scheme B increase roughly linearly. To a certain extent this may be expected as due the smaller demand block sizes, primal residuals tend to be smaller and consequently so are the Lagrangian multiplier updates. On the other hand, the performance of scheme C is roughly independent of the number of clients as the iterations remain constant. This is due to two reasons: 1) independently of the degree of demand disaggregation iterations at the *TSO* level remain constant; 2) the proximal decomposition algorithm is close in principle to a price-based decomposition and thanks to its adaptive penalty update scheme is not significantly affected by the number of subproblems. For small degrees of user disaggregation scheme B performs better as it does not involve the additional round of communications required for the solution of *DSO*^{*} subproblems. For high degrees of decomposition scheme C outperforms scheme B.

- *Case 2 (IEEE RTS 24bus system – contingency)*: In this case a few generators are assumed to be on outage due to a

fault. Demand curtailments are required and as such demand sets the price. This implies strong interactions between *MO* subproblems during convergence which could make the proximal algorithm convergence more difficult. While some increase may be observed in terms of total time for scheme C, convergence time is much better than scheme B. It should be noted that compared to the previous case convergence time is also increased. This is typical behavior of *ADMM* when it has to converge to a particularly high price (*VOLL*).

- *Case 3 (IEEE 300bus system – base data)*: This case differs from the first in that the *TSO* problem is much larger and takes longer to solve. Typically here *DSO*^{*} subproblems are faster to solve, thus the number of *TSO*-level iterations determines the overall convergence time. As may be seen scheme C practically outperforms scheme B in every case.

It should be noted that in all the above test cases the linear increase in iterations for scheme B actually represents a worst case performance, as convergence could be potentially improved through a suitable modification of the *ADMM* penalty factors. However this could be challenging to do for different operating cases in a system without affecting the *TSO*-level iterations. For large systems this is particularly important as a larger number of iterations at that level directly implies an increased convergence time. In addition it should be noted that a more complex and accurate latency model would probably reinforce our conclusions as the communications burden for scheme B is generally higher than that of scheme C.

It is well-known that aggregators are considered to be a fundamental part of the future smart grid [38]. The way these are organized and the algorithms they utilize will have a significant impact on the convergence speed of any decentralized power system operation scheme. Given that the point of delivery of energy in a power system does matter, a collection of microgrids combined with distribution network aggregators could be the natural basis for disaggregated power systems operations in the future. As it is, our results indicate that this is indeed an efficient decomposition structure from a distributed optimization perspective, where increased granularity in demand does not necessarily imply slower convergence.

D. On Decentralized Schemes Convergence

It is a fact that *ADMM* has been proven to converge only for convex problems, whereas the *OPF* problem is a non-linear and non-convex one. As reference [25] points out convergence for non-convex cases cannot be guaranteed. The

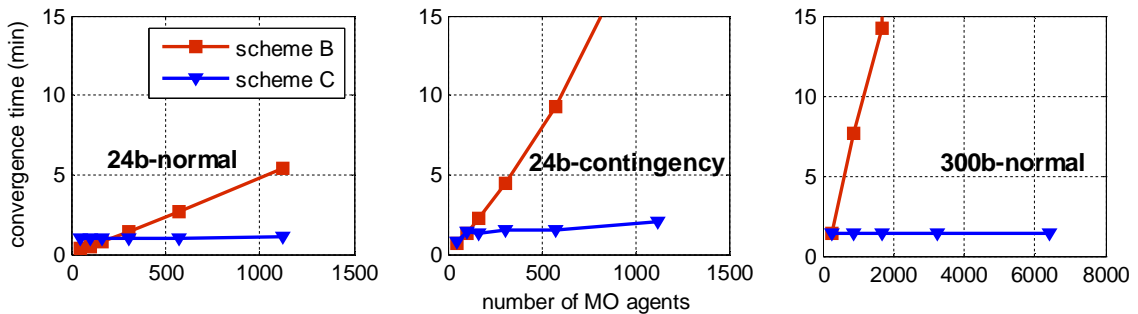


Fig.10. Time to convergence for three different test cases for decomposition schemes B and C. As may be seen scheme C consistently outperforms scheme B as the degree of disaggregation increases.

results presented in Fig.7 are indicative of this fact. On the other hand, for a suitable selection of penalty factors, the method always converged. In terms of optimality, again as [25] points out, the method should be considered as a local optimization method, and as such its performance is dependent on the initial conditions. This however does not imply any worse performance than centralized methods [16]. This was also verified through our simulation results, where the solution of both centralized (interior point method based) and distributed (*ADMM*-based) approaches was the essentially the same. Overall, while a mathematical convergence proof cannot be provided for the general non-convex case, our extensive simulation results indicate that *ADMM* can work reliably for the *OPF* problem, independently of its formulation. Of course, if the recent efforts in convexifying the *OPF* problem (e.g. [39, 40, 41]) find general application then any such convergence issues would be obsolete.

Another question of interest is convergence performance in degenerate cases. As indicated in [42] there are two common types of degeneracy in the *OPF* problem, related to controls and constraints respectively. Control degenerate cases were included in our test set (e.g. load curtailment cases where multiple demand blocks at a single bus were marginal at the same time). This type of degeneracy is resolved through the quadratic terms included in the subproblems and convergence can be ensured. On the other hand, constraint degenerate cases were not involved in our tests. Such cases are difficult to identify but could appear on more complex formulations of the problem with more complex market rules. Relevant investigation could be a direction for future research.

V. CONCLUSIONS

As power systems are becoming ever more granular due to the increased penetration of small renewable generators, distributed storage and electric vehicles, there is a fundamental question to what extent power system operation and control can be decomposed into smaller units. In this paper the answer to that question was investigated within the context of an *OPF* problem, considering both the traditional network decomposition into parallel areas but extended down to the individual bus level, and the decomposition of demand connected to a given bus into individual clients. A parallel implementation of the *ADMM* algorithm has been proposed as a basis for the decomposition of non-convex formulations of the *OPF* problem coupled with an additional proximal decomposition scheme to improve scalability. Extensive simulations have been performed that illustrate the convergence properties of the proposed scheme, and provide insight on the extent that decomposition for an *OPF* problem is feasible.

Overall the presented results indicate that convergence speed is largely dependent on the peculiarity of the particular case to be solved and the way it is decomposed, rather than the number of involved constraints. Slow convergence to high price values (e.g. during load curtailments) and poor propagation of certain constraints (in cases of network congestion) are problems of the method used in this work. Regardless of these facts, the algorithm did eventually converge in every tested

case with sufficient accuracy. With a fast communications infrastructure and a suitably selected partitioning at the transmission level, the proposed scheme does seem able to solve problems fast enough to be of practical value. No financially sensitive information is exchanged between agents, and the scheme is robust to loss of any of its individual components. Furthermore an extension to a multi-period optimization problem is straightforward, with transmission system agents handling independent subproblems for each time period, and other agents handling time-linkage constraints.

It should be emphasized that this work is just a first step towards developing a decentralized structure for energy markets and further research is required. Improving the price update mechanism, extending the algorithm to facilitate the solution of problems with discrete constraints, investigating the effects of security and distribution network constraints, ancillary services and bidding practices, are all key future research targets. In addition, development of accurate communications latency models and investigation of asynchronous algorithms would also be of significant value.

VI. APPENDIX

A. Decomposition Example

In order to clearly illustrate how our decomposition method works we present a simple example for scheme C. Following is the initial problem formulation for the simple network illustrated on Fig.11. For brevity and ease of presentation we use only active power balance equations where the amplitude of all voltage vectors is equal to unity:

$$\min_{\mathbf{P}} \left\{ \begin{array}{l} \mathbf{P} = \text{real}\{\text{diag}(\mathbf{V})(\mathbf{YV})^*\} \\ \sum_{i=1,3} c_i P_i: \mathbf{V} = [V_1, V_2], \mathbf{P} = [P_1, P_2 + P_3] \\ \underline{P}_i \leq P_i \leq \bar{P}_i, i = 1 \dots 3 \end{array} \right\}$$

Now we introduce fictitious buses and nodes (3, 4 and 5), duplicate the corresponding variables, and rewrite the system equations which yields:

$$\min_{\mathbf{P}} \left\{ \begin{array}{l} \mathbf{P}_1 = \text{real}\{\text{diag}(\mathbf{V}_1)(\mathbf{Y}/2 \mathbf{V}_1)^*\} \\ \mathbf{V}_1 = [V_1, V_{3a}], \mathbf{P}_1 = [P_{4b}, P_{3a}] \\ \mathbf{P}_2 = \text{real}\{\text{diag}(\mathbf{V}_2)(\mathbf{Y}/2 \mathbf{V}_2)^*\} \\ \mathbf{V}_2 = [V_2, V_{3b}], \mathbf{P}_2 = [P_{5b}, P_{3b}] \\ \sum_{i=1,3} c_i P_i: \underline{P}_i \leq P_i \leq \bar{P}_i, i = 1, 2, 3 \\ P_1 - P_{4a} = 0 \\ P_2 + P_3 - P_{5a} = 0 \\ \delta_{3a} - \delta_{3b} = 0 \\ P_{ia} + P_{ib} = 0, i = 3, 4, 5 \end{array} \right\} \left\{ \begin{array}{l} h_e(\mathbf{U}) \leq 0 \\ h_c(\mathbf{U}_e) = 0 \end{array} \right.$$

Where $\delta_i = \text{angle}\{V_i\}$. With $\mathbf{z} = [P_{3a}, P_{3b}, \dots, \delta_{3a}, \delta_{3b}]$ based on (14) for *ADMM* iteration k we get two *TSO* subproblems, e.g.:

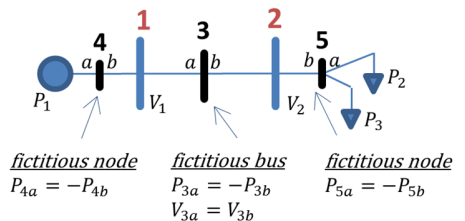


Fig.11. Simple decomposition example.

$$\min_{\substack{P_{3a} \\ P_{4b} \\ \delta_{3a}}} \left\{ \begin{aligned} & -\lambda_{e(1)}^{k-1} P_{3a} - \lambda_{e(4)}^{k-1} P_{4b} - \lambda_{e(7)}^{k-1} \delta_{3a} \\ & + \frac{\rho}{2} ((P_{3a} - \mathbf{z}_{(1)}^{k-1})^2 + (P_{4b} - \mathbf{z}_{(4)}^{k-1})^2 + (\delta_{3a} - \mathbf{z}_{(7)}^{k-1})^2) \end{aligned} \right\}$$

$$\mathbf{P}_1 = \text{real}\{\text{diag}(\mathbf{V}_1)(\mathbf{Y}/2\mathbf{V}_1)^*\}$$

$$\mathbf{V}_1 = [V_1, V_{3a}], \mathbf{P}_1 = [P_{4b}, P_{3a}]$$

In addition we have the subproblem of the generator:

$$\min_{P_1} \left\{ c_1 P_1 - \lambda_{e(3)}^{k-1} P_{4a} + \frac{\rho}{2} (P_{4a} - \mathbf{z}_{(3)}^{k-1})^2 : \frac{P_1}{P_1 - P_{4a}} = 0 \right\}$$

And the DSO* subproblem:

$$\min_{\substack{P_2 \\ P_3 \\ P_{5a}}} \left\{ \begin{aligned} & c_2 P_2 + c_3 P_3 - \lambda_{e(5)}^{k-1} P_{5a} + \frac{\rho}{2} (P_{5a} - \mathbf{z}_{(5)}^{k-1})^2 \\ & \underline{P}_2 \leq P_2 \leq \overline{P}_2 \\ & \underline{P}_3 \leq P_3 \leq \overline{P}_3 \\ & P_2 + P_3 - P_{5a} = 0 \end{aligned} \right\}$$

This subproblem is further decomposed using the APDM algorithm, yielding for iteration l two MO subproblems based on (18), e.g.:

$$\min_{P_2} \left\{ c_2 P_2 - \lambda_{e(5)}^{l-1} P_2 + \frac{\rho}{2} (P_2 - P_2^{l-1})^2 : \underline{P}_2 \leq P_2 \leq \overline{P}_2 \right\}$$

And in addition the DSO price update equation based on (19):

$$\lambda_{e(5)}^l = \lambda_{e(5)}^{k-1} + \rho(P_{5a}^l - \mathbf{z}_{(5)}^{k-1}) \text{ and } P_2^l + P_3^l - P_{5a}^l = 0$$

This completes the description of all possible types of subproblems involved.

B. Multipliers & Auxiliary Variables Updates

Let us assume that $(\lambda_e^k)^T \mathbf{z} = 0 \forall \{\mathbf{z}: \mathbf{h}_c \mathbf{z} = 0\}$. The Lagrangian of (12) is:

$$\mathcal{L} = f_o(\mathbf{U}^{k+1}) + (\lambda_e^k)^T (\mathbf{U}^{k+1} - \mathbf{z}) + \frac{\rho}{2} \|\mathbf{U}^{k+1} - \mathbf{z}\|_2^2 + \boldsymbol{\mu}^T \mathbf{h}_c \mathbf{z}$$

Where $\boldsymbol{\mu}$ are the Lagrange multipliers corresponding to the constraints $\mathbf{h}_c \mathbf{z} = 0$. Removing the terms which are constant and do not affect the optimal \mathbf{z} value, the above simplifies to:

$$\mathcal{L} = \frac{\rho}{2} \|\mathbf{U}^{k+1} - \mathbf{z}\|_2^2 + \boldsymbol{\mu}^T \mathbf{h}_c \mathbf{z}$$

Based on the first order KKT conditions:

$$-\rho(\mathbf{U}^{k+1} - \mathbf{z}^{k+1})^T + \boldsymbol{\mu}^T \mathbf{h}_c = 0 \rightarrow$$

$$\rightarrow \rho(\mathbf{U}^{k+1} - \mathbf{z}^{k+1})^T \mathbf{z} = \boldsymbol{\mu}^T \mathbf{h}_c \mathbf{z} = 0$$

Which is valid for any $\{\mathbf{z}: \mathbf{h}_c \mathbf{z} = 0\}$. From (16) it follows:

$$\lambda_e^{k+1} = \lambda_e^k + \rho(\mathbf{U}^{k+1} - \mathbf{z}^{k+1}) \rightarrow$$

$$\rightarrow (\lambda_e^{k+1})^T \mathbf{z} = (\lambda_e^k)^T \mathbf{z} + \rho(\mathbf{U}^{k+1} - \mathbf{z}^{k+1})^T \mathbf{z} = 0$$

This completes the proof and justifies rewriting (12) as (15).

In order to further clarify the λ_e, \mathbf{z} updates, referring again to Fig.11, let us consider the constraint $P_{3a} + P_{3b} = 0$ and corresponding auxiliary variables $\mathbf{z}_{(1)}, \mathbf{z}_{(2)}$. Let $\lambda_{e(1)}, \lambda_{e(2)}$ be the Lagrange multipliers corresponding to the constraints $\mathbf{z}_{(1)} = P_{3a}, \mathbf{z}_{(2)} = P_{3b}$ and assume their initial values are selected so that $\lambda_{e(1)}^0 = \lambda_{e(2)}^0$. It follows from (12) that $\mathbf{z}_{(1)}^1 = -\mathbf{z}_{(2)}^1 = z^1$ and also:

$$z^1 = \underset{z}{\text{argmin}} \left\{ \lambda_{e(2)}^0 z - \lambda_{e(1)}^0 z + \frac{\rho}{2} (P_{3a}^1 - z)^2 + \frac{\rho}{2} (P_{3b}^1 + z)^2 \right\} \rightarrow$$

$$\rightarrow (P_{3a}^1 - z^1) - (P_{3b}^1 + z^1) = 0 \rightarrow z^1 = (P_{3a}^1 - P_{3b}^1)/2$$

From (13) we then have:

$$\lambda_{e(1)}^1 = \lambda_{e(1)}^0 + \rho(P_{3a}^1 - (P_{3a}^1 - P_{3b}^1)/2) = \lambda_{e(1)}^0 + \frac{\rho}{2}(P_{3a}^1 + P_{3b}^1)$$

$$\lambda_{e(2)}^1 = \lambda_{e(2)}^0 + \rho(P_{3b}^1 + (P_{3a}^1 - P_{3b}^1)/2) = \lambda_{e(2)}^0 + \frac{\rho}{2}(P_{3b}^1 + P_{3a}^1)$$

As may be seen $\lambda_{e(1)}^1 = \lambda_{e(2)}^1$. It should be clear that closed form solutions of (12) are used which implies a negligible computational burden.

VII. ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers whose constructive comments significantly helped to improve this paper.

VIII. REFERENCES

- [1] D. Papadaskalopoulos and G. Strbac, "Decentralized Participation of Flexible Demand in Electricity Markets - Part I: Market Mechanism," *IEEE Trans. Power Systems*, vol. 28, no. 4, pp. 3658-3666, 2013.
- [2] A. J. Conejo, F. D. Galiana, J. M. Arroyo, R. G. Bertrand, C. W. Chua and M. Huneault, "Economic Inefficiencies and Cross-Subsidies in an Auction-Based Electricity Pool," *IEEE Trans. Power Systems*, vol. 18, no. 1, pp. 221-228, 2003.
- [3] M. Roozbehani, M. A. Dahleh and S. Mitter, "Volatility of Power Grids Under Real-Time Pricing," *IEEE Trans. Power Systems*, vol. 27, no. 4, pp. 1926-1940, 2012.
- [4] A. J. Conejo and J. A. Aguado, "Multi Area Coordinated Decentralized DC Optimal Power Flow," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1272-1278, Nov. 1998.
- [5] J. Batut and A. Renaud, "Daily Generation Scheduling Optimization with Transmission Constraints: A New Class of Algorithms," *Trans. Power Systems*, vol. 7, no. 3, pp. 982-989, 1992.
- [6] A. L. Motto, F. D. Galiana, A. J. Conejo and M. Huneault, "On Walrasian Equilibrium for Pool-Based Markets," *IEEE Trans. Power Systems*, vol. 17, no. 3, pp. 774-781, 2002.
- [7] G. Cohen, "Auxiliary Problem Principle and Decomposition of Optimization Problems," *Journal of Optimization Theory and Applications*, vol. 32, no. 3, pp. 277-305, Nov. 1980.
- [8] B. H. Kim and R. Baldick, "Coarse-Grained Distributed Optimal Power Flow," *IEEE Trans. on Power Systems*, vol. 12, no. 2, pp. 932-939, May 1997.
- [9] R. Baldick, B. H. Kim, C. Chase and Y. Luo, "A Fast Distributed Implementation of Optimal Power Flow," *IEEE Trans. Power Systems*, vol. 14, no. 3, pp. 858-864, Aug. 1999.
- [10] J. Contreras, A. Lodi, M. Russo and F. F. Wu, "DistOpt: A Software Framework for Modeling and Evaluating Optimization Problem Solutions in Distributed Environments," *Journal of Parallel and Distributed Computing*, vol. 60, no. 6, pp. 741-763, June 2000.
- [11] A. Lodi and M. Russo, "On the Application of the Auxiliary Problem Principle," *Journal of Optimization Theory and Applications*, vol. 117, no. 2, pp. 377-396, May 2003.
- [12] B. H. Kim, "An Efficient Implementation of Decentralized Optimal Power Flow," *Journal of Electrical Engineering & Technology*, vol. 2, no. 3, pp. 335-341, 2007.
- [13] B. H. Kim and R. Baldick, "A Comparison of Distributed Optimal Power Flow Algorithms," *IEEE Trans. Power Systems*, vol. 15, no. 2, pp. 599-604, May 2000.
- [14] M. Kraning, E. Chu, J. Lavaei and S. Boyd, "Message Passing for Dynamic Network Energy Management," *Foundations and Trends in Optimization*, vol. 1, no. 2, pp. 70-122, 2013.
- [15] A. X. Sun, D. T. Phan and S. Ghosh, "Fully Decentralized AC Optimal Power Flow Algorithms," in *IEEE PES General Meeting*, Vancouver, BC, Canada, 2013.
- [16] T. Erseghe, "Distributed Optimal Power Flow Using ADMM," *IEEE Trans. Power Systems*, vol. 29, no. 5, pp. 2370-2380, 2014.
- [17] A. J. Conejo, F. J. Nogales and F. J. Prieto, "A Decomposition Procedure Based on Approximate Newton Directions," *Mathematical Programming*, vol. 93, no. 3, pp. 495-515, Dec. 2002.
- [18] A. Bakirtzis and P. N. Biskas, "A Decentralized Solution to the DC OPF of Interconnected Power Systems," *IEEE Trans. on Power Systems*, vol. 18, no. 3, pp. 1007-1013, Aug. 2003.
- [19] G. H. Glanzmann and G. Andersson, "Decentralized Optimal Power Flow Control for Overlapping Areas in Power Systems," *IEEE Trans. Power Systems*, vol. 24, no. 1, pp. 327-336, Feb. 2009.
- [20] M. Arnold, S. Knopfli and G. Andersson, "Improvement of OPF Decomposition Methods Applied to Multi Area Power Systems," in

- IEEE PowerTech*, Lausanne, Switzerland, 2007.
- [21] E. Dall'Anese, H. Zhu and G. B. Giannakis, "Distributed Optimal Power Flow for Smart Microgrids," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1464-1475, 2013.
 - [22] E. Loukarakis, J. W. Bialek and C. J. Dent, "A General Framework for Decentralized Energy Trading in Electricity Markets," in *IREP Symposium on Bulk Power System Dynamics and Control*, Rethymnon, Greece, 25-30 August, 2013.
 - [23] N. Gatsis and G. B. Giannakis, "Decomposition Algorithms for Market Clearing With Large-Scale Demand Response," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 1976-1987, 2013.
 - [24] R. D. Christie, B. F. Wollenberg and I. Wangersten, "Transmission Management in the Deregulated Environment," *IEEE Proceedings*, vol. 88, no. 2, pp. 170-195, 2000.
 - [25] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2010.
 - [26] L. Gan, U. Topcu and S. H. Low, "Optimal Decentralized Protocol for Electric Vehicle Charging," *IEEE Trans. Power Systems*, vol. 28, no. 2, pp. 940-951, 2013.
 - [27] A. Y. Ng, M. I. Jordan and Y. Weiss, "On Spectral Clustering: Analysis and an Algorithm," in *Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2001.
 - [28] P. Lagonotte, J. C. Sabonnadiere, J. Y. Leost and J. P. Paul, "Structural Analysis of the Electrical System: Application to Secondary Voltage Control in France," *IEEE Trans. Power Systems*, vol. 4, no. 2, pp. 479-486, 1989.
 - [29] R. J. Sanchez-Garcia, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki and J. W. Bialek, "Hierarchical Spectral Clustering of Power Grids," *IEEE Trans. Power Systems*, 2014.
 - [30] E. Cotilla-Sanchez, P. D. H. Hines, C. Barrows, S. Blumsack and M. Patel, "Multi-Attribute Partitioning of Power Networks Based on Electrical Distance," *IEEE Trans. Power Systems*, vol. 28, no. 4, pp. 4979-4987, 2013.
 - [31] I. Kamwa, A. K. Pradhan, G. Joos and S. R. Samantaray, "Fuzzy Partitioning of a Real Power System for Dynamic Vulnerability Assessment," *IEEE Trans. Power Systems*, vol. 24, no. 3, pp. 1356-1365, 2009.
 - [32] H. Liang, X. Gu and D. Zhao, "Optimization of System Partitioning Schemes for Power System Black-Start Restoration Based on Genetic Algorithms," in *(APPEEC) Power & Energy Engineering Conference*, Chengdu, China, 2010.
 - [33] C. P. Rodriguez and G. J. Anders, "Bidding Strategy Design for Different Types of Electric Power Market Participants," *IEEE Trans. Power Systems*, vol. 19, no. 2, pp. 964-971, 2004.
 - [34] V. P. Gountis and A. G. Bakirtzis, "Bidding Strategies for Electricity Producers in a Competitive Electricity Marketplace," *IEEE Trans. Power Systems*, vol. 19, no. 1, pp. 356-365, 2004.
 - [35] Z. Zhou, F. Zhao and J. Wang, "Agent-Based Electricity Market Simulation With Demand Response From Commercial Buildings," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 580-588, 2011.
 - [36] R. Ma, H. H. Chen, Y. R. Huang and W. Meng, "Smart Grid Communication: Its Challenges and Opportunities," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 36-46, 2013.
 - [37] Y. Yan, Y. Qian and D. Tipper, "A Survey on Smart Grid Communication Infrastructures: Motivations, Requirements and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 5-20, 2013.
 - [38] C. Guille and G. Gross, "A Conceptual Framework for the Vehicle-to-Grid Implementation," *Energy Policy*, vol. 37, no. 11, pp. 4379-4390, 2009.
 - [39] J. Lavaei and S. H. Low, "Zero Duality Gap in Optimal Power Flow Problem," *IEEE Trans. on Power Systems*, vol. 27, no. 1, pp. 92-107, 2012.
 - [40] M. Farivar and S. H. Low, "Branch Flow Model: Relaxations and Convexification - Part I," *IEEE Trans. Power Systems*, vol. 28, no. 3, pp. 2554-2564, 2013.
 - [41] D. K. Molzahn, B. C. Lesieutre and C. L. DeMarco, "A Sufficient Condition for Global Optimality of Solutions to the Optimal Power Flow Problem," *IEEE Trans. Power Systems*, vol. 29, no. 2, pp. 978-979, 2014.
 - [42] B. Stott and O. Alsac, "Optimal Power Flow - Basic Requirements for Real-Life Problems and their Solutions," http://www.ieee.hr/_download/repository/Stott-Alsac-OPF-White-Paper.pdf.

IX. BIOGRAPHIES

Emmanouil Loukarakis (M'12) received the Diploma in Electrical and Computer Engineering from the University of Patras, Greece in 2007, and his M.Sc. Degree from the Technical University of Crete, Greece in 2012. He is currently a postgraduate student at Durham University, UK. His area of interest includes power system economics, dynamics and probabilistic assessment. He is a registered member of the Technical Chamber of Greece and the IEEE Power and Energy Society.

Janusz W. Bialek (F'11) received the M.Eng and Ph.D. degrees from Warsaw University of Technology, Warsaw, Poland, in 1977 and 1981, respectively. In 1981-1989, he was with Warsaw University of Technology, in 1989-2002 with Durham University, and in 2003-2009 with The University of Edinburgh, Edinburgh, U.K. In 2009-2014 he held Chair of Electrical Power and Control at Durham University, and in 2014 he moved to Skolkovo Institute of Science and Technology (Skoltech) in Moscow to lead the Center for Energy Systems. He has co-authored 2 books and over 100 technical papers. His research interests include power system economics, power system dynamics and sustainable energy systems.

Chris J. Dent (M'08) received the B.A. degree in mathematics from Cambridge University, Cambridge, U.K., in 1997, the Ph.D. degree in theoretical physics from Loughborough University, Loughborough, U.K., in 2001, and the M.Sc. degree in operational research from the University of Edinburgh, U.K., in 2006. He is a Research Fellow in the School of Engineering and Computing Sciences, Durham University, U.K. From 2007-2009, he was with the University of Edinburgh. His research interests lie in power system optimization, risk modelling, economics, and renewables integration. He is a Chartered Physicist, an Associate Fellow of the Operational Research Society, and a member of the IET and Cigré.